

## Efeitos de Classe de Medicamentos: Definição e Aplicações Práticas [73]

ISABEL SOARES, ANTÓNIO VAZ CARNEIRO

Serviço de Medicina I, Hospital de Egas Moniz, Lisboa  
Centro de Estudos de Medicina Baseada na Evidência da Faculdade de Medicina de Lisboa, Lisboa

**Rev Port Cardiol 2002;21 (9): 1031-1042**

### RESUMO

O conceito de «efeito de classe» de medicamentos está profundamente enraizado na prática clínica. A utilização de drogas julgadas similares na sua acção – e portanto substituíveis – é muito frequente: são exemplos desta realidade terapêutica a utilização de beta-bloqueantes como antihipertensivos, ou de inibidores da enzima de conversão (IECA) na insuficiência cardíaca congestiva.

No entanto, a definição de classe medicamentosa passa pela existência de três conceitos fundamentais: uma estrutura química semelhante (por exemplo, o anel dihidropiridina em certos bloqueadores dos canais de cálcio), um mecanismo de acção semelhante (os beta-bloqueantes bloqueiam os receptores adrenérgicos) ou um efeito farmacológico semelhante (antihipertensores, antianginosos, etc.).

Neste artigo serão descritos os tipos de evidência em que o cardiologista se pode basear para seleccionar um fármaco específico (dentre vários da mesma classe).

Trata-se portanto de uma abordagem exclusivamente clínica, diferente das que seriam necessárias na selecção de medicamentos por exemplo por uma entidade pagadora (mais interessada em questões de custo-eficácia) ou pela própria indústria farmacêutica (mais interessada em que o medicamento que manufactura seja o mais prescrito possível). Como habitualmente, as recomendações serão baseadas em evidências científicas, cuja hierarquização de importância será detalhadamente apresentada.

### Palavras-Chave

Efeito de classe de medicamentos; Níveis de evidência; Iatrogenia

### ABSTRACT

#### Drug Class Effects: Definitions and Practical Applications

The concept of drug class effect is profoundly rooted in clinical practice. The use of drugs seen as similar in their clinical effects – and therefore interchangeable – is very frequent: two examples of this are the use of beta-blockers in arterial hypertension and angiotensin-converting enzyme (ACE) inhibitors in congestive heart failure.

The definition of drug class effect is based on three concepts: a similar chemical structure (for example, the dihydropyridine ring of some calcium channel blockers), a similar mechanism of action (beta-blockers block adrenoceptors), or similar pharmacological effects (antihipertensives, antianginals, etc.).

In this article we will describe the type of evidence that a cardiologist can use in order to select a specific drug (from within a class). It constitutes a clinical approach, different from the one that might be used by a third party payer (more interested in cost-effectiveness issues) or the pharmaceutical industry (more interested in promoting sales). As usual, the recommendations are based on the strength of scientific evidence.

### Key words

Drug class effect; Levels of evidence; Iatrogenic complications; Evidence-based cardiology

## INTRODUÇÃO

O conceito de «efeito de classe» de medicamentos está profundamente enraizado na prática clínica. A utilização de drogas julgadas similares na sua acção – e portanto substituíveis – é muito frequente: são exemplos desta realidade terapêutica a utilização de beta-bloqueantes como antihipertensivos, ou de inibidores da enzima de conversão (IECA) na insuficiência cardíaca congestiva.

Na realidade, as coisas não são tão simples como parecem, já que a definição de classe medicamentosa passa pela existência de três conceitos fundamentais<sup>(1)</sup>: uma estrutura química semelhante (por exemplo, o anel dihidropiridina em certos bloqueadores dos canais de cálcio – BCC), um mecanismo de acção semelhante (por exemplo, os beta-bloqueantes possuem uma actividade bloqueadora dos receptores adrenérgicos) ou um efeito farmacológico semelhante (por exemplo, antihipertensores ou antianginosos – sejam eles quais forem).

Neste artigo serão descritos os tipos de evidência em que o cardiologista se pode basear para o processo de selecção clínica de um fármaco específico (dentre vários da mesma classe) e baseia-se sobretudo em duas fontes principais<sup>(1, 2)</sup>. Trata-se, portanto, de uma abordagem exclusivamente clínica, diferente das que seriam necessárias na selecção de medicamentos por uma entidade pagadora (mais interessada em questões de custo-eficácia) ou pela própria indústria farmacêutica (mais interessada em que a droga que manufactura seja muito prescrita). Como habitualmente, as recomendações serão baseadas em evidências científicas, cuja hierarquização de importância será detalhadamente apresentada<sup>(3)</sup>.

## CLASSES DE FÁRMACOS

Ao longo deste artigo considerar-se-á como classe de fármacos aquele conjunto de compostos que partilha uma estrutura química e mecanismo de acção semelhantes.

Em virtude destes mecanismos de acção semelhantes, pensa-se que, em geral, os fármacos de uma classe produzem efeitos farmacológicos e resultados clínicos semelhantes (efeitos de classe). É esta assunção que está subjacente às Normas de Orientação Clínica, em que a evidência de estudos que usam um ou mais fármacos dentro de uma classe, é extrapolada para os outros fármacos dessa mesma classe. Assim, como já foi dito, os beta-blo-

## INTRODUCTION

The concept of drug class effect is profoundly rooted in clinical practice. The use of drugs seen as similar in their clinical effects – and therefore interchangeable – is very frequent: two examples of this are the use of beta-blockers in arterial hypertension and angiotensin-converting enzyme (ACE) inhibitors in congestive heart failure.

But things are not as simple as they seem. The definition of drug class effect is based on three concepts: a similar chemical structure (for example, the dihydropyridine ring of some calcium channel blockers), a similar mechanism of action (beta-blockers block adrenoreceptors), or similar pharmacological effects (antihypertensives, antianginals, etc.).

In this article we will describe the type of evidence that a cardiologist can use in order to select a specific drug (from within a class) and is based on two main sources<sup>(1, 2)</sup>. It is therefore a purely clinical approach, different from the one that may have to be used by a third party payer (more interested in cost-effectiveness issues) or the pharmaceutical industry (more interested in promoting sales of its own products). As usual, the recommendations are based on the strength of scientific evidence, a hierarchy of which is presented in detail below<sup>(3)</sup>.

## DRUG CLASSES

In this article, a drug class is defined as the class of compounds which share a similar structure and mechanism of action.

As a result of their similar mechanisms of action, drugs of a particular class produce similar pharmacologic effects and clinical outcomes (“class effects”). This assumption underlies clinical practice guidelines in which evidence from studies using one or more drugs within a class is extrapolated to other drugs of the same class. Thus, as already pointed out, beta-blockers as a class are recommended for patients with acute myocardial infarction, and angiotensin-converting enzyme (ACE) inhibitors for patients with congestive heart failure.

It would seem reasonable to accept *a priori* that drugs within the same class exert similar effects, unless there is clear evidence of important differences. However, this assumption can lead to two significant errors in extrapolation with major clinical consequences. Firstly, if a class of drugs (such as the thiazide diure-

queantes como classe são recomendados para doentes com enfarte agudo do miocárdio, e os inibidores da enzima de conversão da angiotensina (IECAs) para doentes com insuficiência cardíaca congestiva.

Pode parecer razoável aceitar à priori que fármacos dentro da mesma classe exercem efeitos semelhantes, a não ser que haja evidência clara da existência de importantes diferenças. No entanto, esta assunção pode levar a dois importantes erros de extrapolação, com consequências clínicas importantes: em primeiro lugar, quando uma classe de fármacos (como os diuréticos tiazídicos) produzem efeitos farmacológicos semelhantes (redução da tensão arterial) e efeitos clínicos semelhantes (redução do número de acidentes vasculares cerebrais) pode-se ser levado a considerar que uma segunda classe de fármacos (por exemplo, bloqueadores dos canais de cálcio), que produzem os mesmos efeitos farmacológicos, também deverão produzir o mesmo benefício clínico. Na ausência de ensaios aleatorizados e controlados (EACs) que verifiquem este passo final, este tipo de extrapolação pode ser errada. Por exemplo, certos bloqueadores dos canais de cálcio têm efeitos desfavoráveis sobre a mortalidade total<sup>(4)</sup>. Por outro lado, mesmo no interior de uma classe, diferentes fármacos podem ter diferentes efeitos fisiológicos, para além do mecanismo de acção que serviu para os classificar como pertencendo à mesma classe. Em consequência, pode ser pouco rigoroso extrapolar os resultados clínicos que foram demonstrados em ensaios aleatorizados de um fármaco numa classe, para outro fármaco da mesma classe que não tenha sido sujeito a ensaios semelhantes. Por exemplo, há autores que consideram que, embora todas as estatinas actuem sobre a HMG-coenzima A redutase, elas podem exercer diferentes efeitos não lipídicos sobre o processo aterotrombótico, que podem influenciar a sua eficácia clínica.

Em termos técnico-científicos, como se devem analisar as propriedades terapêuticas de cada medicamento? O efeito de tratamento absoluto causado por um fármaco, definido pela redução de risco absoluta (RRA) ou número necessário tratar (NNT), é influenciado pelo risco basal ou taxa de eventos do grupo de controlo (TEC) dos doentes em que é usado. Assim, de grupo para grupo de doentes, pode haver uma considerável variação da RRA. Já no que diz respeito ao efeito de tratamento relativo de um fármaco (definido pela redução de

tics) produce similar pharmacological effects (lowering of blood pressure) and similar clinical effects (reduction in occurrence of stroke), it might be assumed that a second class of drugs (for example, the calcium-channel blockers) that produce the same pharmacological effects will also produce the same clinical benefit. In the absence of randomized trials that verify this final step, such extrapolation may be in error. For example, some calcium channel blockers have unfavourable effects on total mortality<sup>(4)</sup>. Secondly, even within the same class, individual drugs may have different physiologic effects, other than the mechanism of action which defined them as belonging to the same class. It may therefore be inaccurate to extrapolate the clinical outcomes shown in randomized trials of one drug in a particular class to another member of that class which has not been subjected to similar trials. For example, some authors have argued that, although all the statins act on HMG-coenzyme A reductase, they may have different non-lipid effects on the atherothrombotic process which may influence their clinical efficacy.

In technical and scientific terms, how should we analyze the therapeutic properties of a given drug? The absolute treatment effect produced by a drug, defined by the absolute risk reduction (ARR) or number needed to treat (NNT), is influenced by the baseline risk or control event rate (CER) of the patients in whom it is used. Thus there may be considerable variation in ARR among different patient groups. With regard to the relative treatment effect of a drug (defined by the relative risk reduction or RRR), this is often (but not always) similar, irrespective of the baseline risk of the trial participants. Accordingly, if two drugs are tested in separate placebo-controlled trials, only proportional effects can be compared, that is, the RRR seen with each drug (and even then only by assuming constant RRR with different CERs)<sup>(5,6)</sup>.

A class effect is considered to exist when drugs with similar mechanisms of action generate RRRs or odds ratios (ORs)<sup>(1)</sup> that are similar in direction and magnitude. For example, the Collaborative Group on ACE Inhibitor Trials<sup>(7)</sup> suggested that there is a class effect for ACE inhibitors in patients with symptomatic heart failure despite the fact that the OR point estimates for effects on total mortality ranged between 0.14 (95 % CI: 0 to 7.6) for perindopril and 0.78 (95 % CI: 0.67 to 0.91)

risco relativo – RRR), verifica-se que muitas vezes é semelhante – embora nem sempre – e independentemente do risco basal dos participantes no ensaio. É por estas razões que quando dois fármacos são testados em ensaios separados, controlados por placebo, só se podem comparar os efeitos proporcionais, ou seja, a RRR observada com cada fármaco (e, mesmo assim, admitindo-se uma RRR constante sob diferentes TECs)<sup>(5, 6)</sup>.

Pode-se considerar que existe um efeito de classe quando fármacos com um mecanismo de acção semelhante geram RRRs ou *odds ratios* (ORs)<sup>1</sup> semelhantes em direcção e amplitude. Por exemplo, o *Collaborative Group on ACE Inhibitor Trial*<sup>(7)</sup> sugeriu que há um efeito de classe dos IECAs em doentes com insuficiência cardíaca congestiva sintomática, apesar do facto das estimativas pontuais das ORs para efeitos sobre a mortalidade total terem variado entre 0,14 (IC a 95%: 0 a 7,6) para o perindopril e 0,78 (IC a 95%: 0,67 a 0,91) para o enalapril. A confiança neste efeito de classe provém do reconhecimento de que a OR global em 32 ensaios era de 0,77 (IC a 95%: 0,67 a 0,88), para além de que havia sobreposição do intervalo de confiança para cada um dos IECAs e não foi encontrada heterogeneidade estatística entre ensaios de diferentes agentes.

No entanto, há riscos em assumir numa forma generalizada um efeito de classe de fármacos, porque nem sempre as coisas se passam da mesma maneira. Por exemplo, o ensaio SWORD sugeriu um aumento na mortalidade com o uso de d-sotalol em doentes com enfarte agudo do miocárdio que desenvolvem insuficiência cardíaca congestiva, ao contrário do que acontece com os outros  $\beta$ -bloqueantes<sup>(8)</sup>.

## NÍVEIS DE EVIDÊNCIA

Quando temos de tomar decisões sobre a eficácia clínica e segurança comparativas de fármacos dentro de uma classe, é útil aplicar uma hierarquia de evidência, para conseguir otimizar o processo de selecção e evitar extrapolações erróneas. O *Quadro 1* apresenta um

for enalapril. Confidence in this class effect stems from the recognition that the overall OR in 32 trials was 0.77 (95% CI: 0.67 to 0.88), plus the fact that the confidence intervals for each of the ACE inhibitors overlapped, and that there was no statistical heterogeneity between trials of different agents.

However, there are risks in assuming too readily that a class effect exists, because the outcomes are not always the same. For example, the SWORD trial suggested an increase in mortality with d-sotalol in patients with acute myocardial infarction who develop heart failure, in contrast to what is found with other beta-blockers<sup>(8)</sup>.

## LEVELS OF EVIDENCE

When we need to make decisions on the comparative clinical efficacy and safety of drugs within a given class, it is useful to apply a hierarchy of evidence to optimize the process of drug selection and avoid erroneous extrapolation. *Table 1* presents a scheme of levels of evidence to be used when comparing one drug against others within the same class<sup>(2)</sup>. Such a comparison should be part of any systematic review of all the relevant evidence on the effects of a particular treatment. The key question to be answered by a meta-analysis of class effects is whether the differences between individual drugs explain the variations in outcomes found in trials or within trials that compare different drugs.

Level 1 studies are randomized trials in which direct comparisons are made between the drug of interest and other drugs of the same class – and not with placebo – with regard to their effects on clinically important outcomes. Although such RCTs that compare drugs produce the strongest evidence, there are, nevertheless, certain points to take into consideration:

- Firstly, it should have been demonstrated previously that at least one of the drugs has a clinically important impact versus placebo in trials carried out in a similar population to that of the current trial.

<sup>1</sup> Traduzido por razão de probabilidades o *odds ratio* define-se da seguinte maneira: *odd* é uma relação entre a probabilidade da ocorrência com a de não-ocorrência de um determinado evento, isto é, uma relação entre a probabilidade que algo seja alguma coisa com a probabilidade que não seja nada. Por exemplo, se de 100 fumadores, 80 desenvolvem uma tosse crónica e 20 não, o *odd* de aparecimento da tosse neste grupo é de 80:20, ou seja, 4; em contraste, se de 100 não fumadores, 5 desenvolvem uma tosse crónica e 95 não, o *odds ratio* de aparecimento da tosse neste grupo é de 5/95 ou seja,  $\pm 0,05$ . O *odds ratio* é a razão entre dois *odds* assim descritos. (4: 0,05=80)

<sup>1</sup> “Odds ratio” is defined as follows: “odds” are the ratio between the probability of a particular event occurring or not occurring, that is the ratio between the probability that something exists and the probability that it does not. For example, if out of 100 smokers 80 develop a chronic cough and 20 do not, the odds of a cough appearing in this group is 80:20, or 4; by contrast, if out of 100 non-smokers, 5 develop a chronic cough and 95 do not, the odds of a appearing in this group are 5:95 or  $\pm 0.05$ . The odds ratio is the ratio between two such odds. (4: 0.05=80)

esquema de níveis de evidência a usar na comparação de um fármaco com outros da mesma classe <sup>(2)</sup>. Esta comparação deve fazer parte de qualquer revisão sistemática de toda a evidência relevante sobre os efeitos de um tratamento. A questão chave a ser esclarecida por uma metanálise no que diz respeito a efeitos de classe é se os fármacos individuais explicam a variabilidade de resultados através dos ensaios – ou dentro dos ensaios que comparam diferentes fármacos.

Os estudos de nível 1 compreendem os ensaios aleatorizados em que são feitas comparações directas entre o fármaco de interesse e outros fármacos da mesma classe – e não com placebos – em relação aos seus efeitos sobre resultados importantes. Apesar destes EACs comparativos de fármacos gerarem a mais forte evidência, há, no entanto, mesmo assim, alguns aspectos a considerar:

- Secondly, the choice of the appropriate dose for each drug is a complicated issue, in that it affects the outcomes and safety profiles for both drugs.
- Finally, one must carefully consider the trial size and methodology before concluding equivalence of the two drugs. Equivalence trials <sup>(9)</sup> require much larger sample sizes than standard trials and any lack of rigor in trial conduct or patient compliance is enough to mask any real differences between the drugs.

The choice of clinically important outcomes for level 1 studies will vary depending on the target intervention. In the case of therapies designed to prevent atherosclerosis (such as statins), clinically important outcomes will include long-term efficacy data on events such as myocardial infarction, stroke, and all-cause mortal-

#### Quadro 1

#### Níveis de evidência para comparação da eficácia de fármacos dentro de uma mesma classe

Nível	Comparação	Doentes do estudo	Resultados	Ameaças à validade
1	Dentro de um EAC fazendo comparação directa entre fármacos	Recrutados da mesma população (por definição)	Resultados importantes para o doente •	Aleatorização não ocultada Ausência de ocultação Follow-up incompleto
2	Dentro de um EAC em paralelo	Idênticos (por definição)	Resultados substitutos validados †	As de nível 1 Validade do resultado substituto para resultados clinicamente importantes
2	Entre EACs de diferentes fármacos <i>versus</i> placebo	Semelhantes ou diferentes (em estado de doença e de factores de risco)	Resultados importantes para o doente ou resultados substitutos validados	As de nível 1 Diferenças entre os ensaios em: • Características dos doentes • Qualidade metodológica (“blinding” adequado, aleatorização ocultada, etc) • Definições de resultados • Taxas de aderência
3	Entre análises de subgrupo de EACs de diferentes fármacos <i>vs</i> placebo	Semelhantes ou diferentes	Resultados importantes para os doentes ou resultados substitutos	As de nível 1 e 2 Múltiplas comparações, após colheita de dados ad hoc Subgrupos com poder limitado Classificação incorrecta em subgrupos
3	Entre EACs de diferentes fármacos <i>vs</i> placebo	Semelhantes ou diferentes	Resultados substitutos não validados	Os resultados substitutos podem não capturar todos os efeitos (benéficos ou prejudiciais) de um agente terapêutico
4	Entre estudos não aleatorizados (estudos observacionais e investigação de bases de dados administrativas)	Semelhantes ou diferentes	Resultados importantes para os doentes	Confundimento por indicação, aderência, ou calendário Confundidores desconhecidos/não avaliados Erro de medição Para investigação de resultados: bases de dados limitadas, sistemas de codificação não adequados para investigação

• Resultados importantes para o doente referem-se a dados de eficácia a longo prazo e os marcadores específicos dependem da condição em tratamento. No caso das estatinas, que são usadas para prevenir e tratar doença aterosclerótica, os resultados importantes para os doentes incluem mortalidade de todas as causas, enfarte do miocárdio e acidente vascular cerebral.

† Resultados substitutos são considerados validados apenas quando a relação entre o resultado substituto e os resultados importantes para o doente tiver sido firmemente estabelecida em EACs a longo prazo.

Quadro adaptado de <sup>(2)</sup>.

Table 1

**Levels of evidence for comparing the efficacy of drugs within the same class**

Level	Comparison	Study patients	Outcomes	Threats to validity
1	Within a “head-to-head” randomized trial (RCT)	Recruited from the same population (by definition)	Clinically important outcomes *	Randomization not concealed No blinding Incomplete follow-up
2	Within a “head-to-head” RCT	Identical (by definition)	Validated surrogate outcomes †	As for level 1: Validity of surrogate outcome for clinically important outcomes
2	Between RCTs of different drugs vs. placebo	Similar or different (in disease and risk factor status)	Clinically important outcomes or validated surrogate outcomes	As for level 1 Differences between trials in: • patient characteristics • methodologic quality (adequacy of blinding, concealment of randomization, etc.) • endpoint definitions • compliance rates
3	Between subgroup analyses from RCTs of different drugs vs. placebo	Similar or different	Clinically important outcomes or validated surrogate outcomes	As for level 1 and 2 • multiple comparisons, post-hoc data dredging • under-powered subgroups • misclassification into subgroups
3	Between RCTs of different drugs vs. placebo	Similar or different	Unvalidated surrogate outcomes	Surrogate outcomes may not capture all the effects (beneficial or harmful) of a therapeutic agent
4	Between non-randomized studies (observational studies and administrative database research)	Similar or different	Clinically important outcomes	Confounding by indication, compliance, and calendar time Unknown/unmeasured confounders Measurement error For outcomes research: limited databases, coding systems not suitable for research

\* “Clinically important outcomes” refer to long-term efficacy data and the particular endpoints depend on the condition being treated. In the case of statins, which are used to prevent and treat atherosclerosis, clinically important outcomes include all-cause mortality, myocardial infarction, and stroke.

† “Surrogate outcomes” are considered validated only when the relation between the surrogate outcome and clinically important outcomes has been firmly established in long-term RCTs.

Table adapted from <sup>(9)</sup>.

- Em primeiro lugar, é necessário que tenha sido demonstrado previamente que pelo menos um dos fármacos teve um impacto clinicamente importante versus placebo, em ensaios realizados numa população que é semelhante à do corrente ensaio.
- Em segundo lugar, a escolha da dose apropriada para cada fármaco é um problema complexo na medida em que afecta os resultados e os perfis de segurança de ambos os fármacos.
- Por último, deve-se considerar cuidadosamente o tamanho do ensaio e a sua metodologia antes de concluir pela equivalência dos dois fármacos. Os ensaios de equivalência <sup>(9)</sup> requerem tamanhos de amostra muito superiores à dos ensaios standards, e uma fraca exigência na condução do ensaio e na aderência dos doentes é suficiente para mascarar qualquer diferença real entre os fármacos.

ity. On the other hand, for interventions designed to treat symptomatic diseases, clinically important outcomes would include symptom scores and other measures of quality of life.

Level 2 studies include:

- RCTs in which direct comparisons are made between the drug of interest and other drugs of the same class regarding their effects on validated surrogate outcomes;
- comparisons between two or more trials in which active ingredients are tested against placebo regarding clinically important outcomes or validated surrogate outcomes.

Direct comparisons regarding surrogate outcomes can only be accepted as level 2 evidence if at least one of the drugs under comparison has previously been shown to be effective in long-term trials with clinically important outcomes. For example, although cohort stud-



Os resultados considerados importantes para estudos nível 1 variam conforme a intervenção alvo: no caso de terapêuticas destinadas a prevenir arterosclerose (como estatinas, por ex.), resultados importantes incluem, por exemplo, dados de eficácia a longo prazo sobre eventos como enfarte agudo do miocárdio, acidente vascular cerebral e mortalidade de todas as causas. Por outro lado, no caso de intervenções destinadas a tratar doenças sintomáticas, os resultados importantes consistem em scores de sintomas e outras medidas de qualidade de vida.

Os estudos de nível 2 compreendem:

- EACs em que são feitas comparações directas do fármaco de interesse com outros fármacos da mesma classe em relação aos seus efeitos sobre resultados substitutos (*surrogate outcomes*) validados;
- Comparações entre dois ou mais ensaios em que são testados agentes activos versus placebo em relação a resultados clinicamente importantes ou resultados substitutos validados.

Só é legítimo aceitar comparações directas relativamente a resultados substitutos como evidência nível 2 se tiver sido provado previamente que pelo menos um dos fármacos em comparação foi eficaz em ensaios a longo prazo com resultados importantes para os doentes. Por exemplo, embora estudos de *coorte* e EACs com agentes hipolipidemiantes suportassem a hipótese de que a redução dos níveis de colesterol LDL baixa o risco de doença cardíaca aterosclerótica, só com a publicação de ensaios de estatinas em larga escala ligando de forma consistente reduções no colesterol LDL a reduções na morbidez e mortalidade, é que foi possível gerar-se a confiança suficiente para aceitar o resultado substituto de baixar o nível de colesterol LDL como um reflexo de resultados importantes para os doentes<sup>(10)</sup>.

Por outro lado, embora formalmente se possa considerar que um ensaio comparando quatro estatinas relativamente aos seus efeitos sobre o colesterol LDL, HDL e triglicéridos, durante um período de oito semanas, seja evidência nível dois, é também importante entrar em conta com o tamanho e duração dos ensaios no processo de decisão.

Entre as várias estatinas é possível fazer ainda outro tipo de comparações de nível 2. É possível, por exemplo, comparar os efeitos de tratamento observados com simvastatina e pra-

ies and RCTs with lipid-lowering agents support the hypothesis that reducing LDL cholesterol levels reduces the risk of atherosclerotic heart disease, it was only with the publication of large-scale statin trials consistently linking reduction in LDL cholesterol to reduced morbidity and mortality that it was possible to feel confident that the surrogate outcome of lowered LDL cholesterol was a reflection of clinically important outcomes<sup>(10)</sup>.

However, although a trial comparing four statins regarding their effects on LDL and HDL cholesterol and triglycerides over a period of eight weeks can theoretically be considered level 2 evidence, it is also important to take into account the size and duration of the trials in the decision-making process.

It is possible to make another type of level 2 comparison between the various statins. One can, for example, compare the treatment effects seen with simvastatin versus pravastatin in the 4S and LIPID secondary prevention trials, in which each of the drugs is tested separately versus placebo. While consistency of effects in such comparisons is strong evidence for the presence of a class effect, they are less useful in determining whether one drug is more efficacious than another, since the advantages of randomization are lost and the comparison is essentially between two or more cohorts.

For example, it is possible that the populations in the trials under comparison may have different baseline risks and that this will lead to different responses to intervention. This is what would occur if one compared a statin used in a primary prevention trial with another used in a secondary prevention trial. The RRR may be greater in secondary prevention trials, irrespective of which statin is used. There is thus the risk of attributing a greater effect to the drug used in the secondary prevention trial, when in fact the difference has nothing to do with the drug itself. Comparisons such as these, between trials with subjects of varying risk, would only be valid if the RRR is independent of baseline risk, which has been questioned for the statins.

It is theoretically possible to compare the efficacy of two drugs tested in separate placebo-controlled trials. An indirect assessment of the relation between drugs A and B can be obtained by comparing the OR or RR from studies of drug A versus placebo (p) with those from studies comparing drug B versus placebo:

vastatina através dos ensaios de prevenção secundária 4S e LIPID em que cada um destes fármacos é testado separadamente versus placebo. A existência de consistência de efeitos neste tipo de comparações é forte evidência para a presença de um efeito de classe, no entanto, a sua utilidade é menor quando se trata de determinar se um fármaco é mais eficaz do que outro, visto que as vantagens da aleatorização são perdidas e a comparação se faz essencialmente entre dois ou mais *coortes*.

Existe a possibilidade, por exemplo, de que as populações usadas nos ensaios em comparação tenham diferentes riscos de base e que isso implique diferentes respostas à intervenção. É o que acontece se se comparar uma estatina usada num ensaio de prevenção primária com outra estatina usada num ensaio de prevenção secundária. A RRR pode ser maior em ensaios de prevenção secundária, independentemente de qual é a estatina usada. Corre-se, assim, o risco de atribuir um efeito maior ao fármaco usado em ensaios de prevenção secundária, quando a diferença não tem nada a ver com o fármaco. Comparações como esta, entre estudos de doentes com risco variável, seriam válidas só se a RRR fosse independente do risco basal, o que tem sido questionado para as estatinas.

É teoricamente possível comparar a eficácia de dois fármacos testados em ensaios controlados por placebo separados. A avaliação indirecta da relação entre os fármacos A e B pode ser obtida comparando a OR ou RR de estudos do fármaco A *versus* placebo (p) com as de estudos do fármaco B *versus* placebo:  $OR(A \text{ vs } B) = OR(A \text{ vs } p) / OR(B \text{ vs } p)$ . Contudo, esta comparação só é possível se se assumir que nenhum dos vieses potenciais mencionados no *Quadro 1* está presente e que o efeito de tratamento é consistente através de diferentes subgrupos de doentes. O que se tem verificado é que estas comparações indirectas fornecem avaliações do tamanho do efeito substancialmente diferentes de comparações directas entre dois fármacos. Assim, a força da inferência de comparações indirectas é limitada.

Os estudos nível 3 e nível 4 têm numerosos inconvenientes e devem ser encarados de preferência como exercícios de geração de hipóteses.

Os estudos de nível 3 compreendem:

- Comparações baseadas em análise de subgrupos de diferentes ensaios controlados por placebo;

$OR A \text{ vs. } B = OR A \text{ vs. } p / OR B \text{ vs. } p$ . However, this comparison is only possible if one assumes that none of the potential biases listed in Table I are operative and that the treatment effect is consistent across different patient subgroups. It has been found that these indirect comparisons provide estimates of effect size that are substantially different from direct comparisons between two drugs. Thus, the strength of inference from indirect comparisons is limited.

Level 3 and 4 studies have numerous disadvantages, and are best viewed as exercises in hypothesis generation.

Level 3 studies include:

- comparisons between subgroups from different placebo-controlled trials;
- comparisons between placebo-controlled trials in which outcomes are restricted to unvalidated surrogate endpoints.

In addition to the biases that affect higher level studies, comparisons based on subgroup analysis are potentially flawed. Experience shows that many initial subgroup conclusions, especially when they are the result of “data-dredging”, are later proved wrong. Level 3 evidence also includes the use of surrogate endpoints which, although they may be on a recognized pathogenic pathway from mechanisms of action to important clinical outcomes, have not been validated in long-term randomized controlled trials.

Level 4 studies are comparisons based on non-randomized evidence. Examples of non-randomized evidence include cohort or case control studies, modeling studies<sup>2</sup> and outcomes research using administrative databases. Such analyses should be viewed as exercises in hypothesis generation.

Some authors have suggested that pravastatin may have additional anti-inflammatory properties to other statins, which could lead to a greater reduction in coronary events than could be attributed to falls in lipid levels. An example of level 4 evidence is a recent re-analysis of the WOSCOPS database which investigated this hypothesis. In this study, the observed coronary event rates in pravastatin-treated patients were compared to those predicted from the Framingham coronary risk equation. Using the combination of risk factors and mean cholesterol levels during treatment observed in the trial to determine the expected event rate, the analysis found a greater reduction in events



- Comparações entre ensaios controlados por placebo em que os resultados são limitados a marcadores substitutos não validados.

Além dos vieses que afectam estudos de nível mais elevado, as comparações baseadas em análises de subgrupo são potencialmente errôneas. A experiência mostra que muitas das conclusões iniciais de subgrupos, especialmente quando são o resultado de sobre-análise de dados, se vêm a provar erradas. Incluído na evidência nível 3 está também o uso de marcadores substitutos que, embora fiquem na sequência de uma via patogénica reconhecida de mecanismos de acção para resultados clínicos importantes, não têm sido validados em ensaios aleatorizados e controlados a longo prazo.

Os estudos de nível 4 consistem em comparações utilizando evidência não aleatorizada. Exemplos de evidência não aleatorizada incluem estudos de coorte ou de caso-controlo, *modeling studies*<sup>2</sup> ou investigação de resultados utilizando bases de dados administrativas. Este tipo de análises deve ser encarado como exercícios de geração de hipóteses.

Alguns autores sugerem que a pravastatina pode ter propriedades anti-inflamatórias para além das outras estatinas, o que pode levar a uma redução nos eventos coronários superior à que pode ser atribuída à descida dos lípidos. Um exemplo de evidência de nível 4 é uma reanálise recente da base de dados WOSCOPS, que explorou esta hipótese. Neste estudo os investigadores compararam as taxas de eventos coronários observadas em doentes tratados com pravastatina com as previsíveis a partir da equação de risco de eventos coronários de Framingham. Foram usados o conjunto dos factores de risco e níveis de colesterol médios durante o tratamento observados no ensaio para determinar a taxa de eventos esperada, descobrindo uma maior redução nos eventos do que seria de esperar da diferença dos níveis de colesterol entre os dois grupos. Os autores inferiram que a pravastatina pode ter maior eficácia do que as outras estatinas<sup>(1)</sup>.

## **DIMENSÃO DA EVIDÊNCIA, SEGURANÇA, ADERÊNCIA E CUSTO DE FÁRMACOS**

Embora evidência de nível 1 seja a mais adequada para demonstrar que um grupo de

than would be expected from the difference in cholesterol levels between the two groups. The authors inferred that pravastatin may be more efficacious than the other statins<sup>(1)</sup>.

## **AMOUNT OF EVIDENCE, SAFETY, COMPLIANCE AND COST OF DRUGS**

While level 1 evidence is the most appropriate for demonstrating that a drug group has a class effect, or for proving the superiority of one drug over another within the same class, we must recognize that such evidence is rare and unlikely to ever be available for many drug classes, owing to the difficulty of funding and conducting large-scale trials. In this situation, the amount of level 2 evidence becomes important. For instance, one would have more confidence in concluding a class effect if there were several placebo-controlled trials demonstrating that various drugs from the same class had similar treatment effects. The size and duration of the studies are also crucial to the decision-making process.

It is known that drugs within the same class can have different safety profiles (for example, ticlopidine causes more neutropenia than clopidogrel; phenylbutazone causes agranulocytosis while other non-steroidal anti-inflammatory drugs do not). *Table II* presents a set of levels of evidence for determining the safety of drugs within the same class<sup>(2)</sup>.

The first tests of a drug in humans (phase I studies) are designed to determine the maximum tolerated dose; clinical trials (phase II and III trials) are designed to assess a drug's efficacy. As such, none of these studies use sufficiently large samples to detect rare adverse effects. The "inverse rule of 3" states that to have a 95% probability of seeing an adverse reaction to a drug that occurs in every x patients, it is necessary to follow 3x patients. Given the size and duration of most clinical trials, adverse effects that occur in less than 1 in every 1000 participants, or take more than 6 months to appear, will usually go undetected.

Besides being generally too small to detect rare adverse effects, RCTA have other limitations as a method of assessing drug toxicity. In these trials, a high risk of drug toxicity is often a patient exclusion criterion, and more attention is usually given to monitoring the reduc-

<sup>2</sup> estudos de modelação de regras de predição clínica, por exemplo, baseadas em ensaios clínicos que definem as variáveis preditoras.

<sup>2</sup> modeling studies that model guidelines of clinical prediction, such as those based on clinical trials that define predictor variables.

fármacos exerce um efeito de classe, ou para provar a superioridade de um fármaco dentro de uma classe, tem que se reconhecer que tal evidência raramente existe e é improvável que alguma vez venha a existir para muitas classes de fármacos devido a dificuldades em financiar e conduzir grandes ensaios.

Neste caso, a quantidade de evidência de nível 2 torna-se importante, sendo, por exemplo, possível concluir com mais confiança pela existência de um efeito de classe se houver vários ensaios controlados por placebo a demonstrar que vários fármacos da mesma classe têm efeitos de tratamento semelhantes. O tamanho e a duração dos estudos também são essenciais para o processo de decisão.

É um facto conhecido que, dentro da mesma classe, os fármacos podem ter perfis de segurança diferentes (por exemplo, ticlopidina causa mais neutropénia do que clopidogrel; fenilbutazona causa agranulocitose, ao contrário do que acontece com outros anti-inflamatórios não esteróides). No *Quadro II* é mencionado um conjunto de níveis de evidência para determinar a segurança de fármacos dentro da mesma classe<sup>(2)</sup>.

Os primeiros testes de um fármaco em seres humanos (estudos fase I) têm por objectivo determinar a dose máxima tolerada; os ensaios clínicos (estudos fase II e fase III) têm por objectivo avaliar a eficácia de um fármaco. Como tal, nenhum destes estudos utiliza amostras de tamanho adequado para detectar efeitos adversos raros. A «regra inversa de 3» estipula que para ter 95 % de certeza de ver uma reacção adversa a um fármaco que ocorre em cada X

tion of negative target outcomes than to monitoring toxicity. However, when toxic effects are strictly monitored, RCTA remain the best method of detecting real differences in adverse effects between drugs, such as the different rates of intracranial bleeding found with different thrombolytic agents. In the absence of data on toxicity from high-quality RCTA, pre-marketing safety data must be considered preliminary, and large phase IV studies (or systematic analysis of post-marketing surveillance data) are necessary to confirm the safety of new drugs.

#### *Quadro II*

#### **Níveis de evidência para comparação da segurança de fármacos dentro da mesma classe**

Nível	Propriedades	Vantagens	Ameaças à validade
1	EACs	Único desenho que permite a detecção de efeitos adversos quando o efeito adverso é semelhante ao evento que o tratamento está a tentar impedir	Falta de poder estatístico (subpotenciado para detectar efeitos adversos)
2	<i>Coorte</i>	Colheita de dados prospectivos, coorte definido	Depende criticamente do rigor do seguimento, classificação e avaliação
3	Caso-controlo	Pouco dispendioso e de execução rápida	Viés de selecção e de lembrança; a relação temporal pode não ser clara
4	Vigilância pós-marketing	Se suficientemente ampla, pode detectar efeitos adversos raros mas importantes	Ausência de grupo de controlo; depende criticamente do rigor do seguimento, classificação e avaliação
5	Série de casos	Pouco dispendioso e de execução rápida	Amostra de pequena dimensão; viés de selecção; ausência de grupo de controlo
6	Casos clínicos	Pouco dispendioso e de execução rápida	Amostra de pequena dimensão; viés de selecção; ausência de grupo de controlo

Table II

**Levels of evidence for comparing the safety of drugs within the same class**

Level	Properties	Advantages	Threats to validity
1	RCT	The only design that enables detection of adverse effects when the adverse effect is similar to the event which the treatment is intended to prevent	Underpowered for detecting adverse effects
2	Cohort	Prospective data collection, defined cohort	Depends critically on follow-up, classification and assessment
3	Case-control	Inexpensive and rapid to perform	Bias from selection and recall; the temporal relationship may not be clear
4	Post-marketing surveillance	If sufficiently large, can detect rare but important adverse effects	No control group; depends critically on follow-up, classification and assessment
5	Case series	Inexpensive and rapid to perform	Small sample size; selection bias; no control group
6	Case studies	Inexpensive and rapid to perform	Small sample size; selection bias; no control group

Table adapted from <sup>(2)</sup>.

doentes é necessário seguir 3X doentes. Dado o tamanho e a duração da maior parte dos ensaios clínicos, os efeitos adversos que ocorrem em menos de 1 em cada 1000 participantes ou levam mais de 6 meses a aparecer, geralmente permanecem não detectados.

Para além de serem, em geral, demasiado pequenos para detectar efeitos adversos raros, os EACs têm ainda outras limitações, como metodologia para avaliar a toxicidade de fármacos. Nestes ensaios, a presença de alto risco de toxicidade de fármacos é muitas vezes um critério de exclusão de doentes, e, em regra, é dada mais atenção à monitorização da redução dos resultados alvo negativos do que à monitorização da toxicidade. No entanto, quando a monitorização dos efeitos tóxicos é rigorosa, os EACs constituem o método mais forte para detectar diferenças reais entre fármacos relativamente a efeitos adversos, como acontece, por exemplo, com as diferentes percentagens de hemorragia intracraniana que se verificam entre os diferentes agentes trombolíticos. Na ausência de dados de EACs de qualidade a respeito da toxicidade, os dados de segurança pré-marketing devem ser considerados preliminares, sendo necessários grandes estudos fase IV (ou análise sistemática de dados de vigilância pós-marketing – farmacovigilância) para confirmar a segurança de novos fármacos.

Pedido de separatas para:

Address for reprints:

A. VAZ CARNEIRO

Centro de Estudos de Medicina Baseada na Evidência

Faculdade de Medicina de Lisboa

1649-035 LISBOA

## BIBLIOGRAFIA / REFERENCES

1. McAlister FA, Laupacis A, Wells GA, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. XIX. Applying clinical trial results. B. Guidelines for determining whether a drug is exerting (more than) a class effect. JAMA 1999;282:1371-7.
2. Guyatt GH, Rennie D, editors. User's Guides to the Medical Literature. A Manual for Evidence-Based Clinical Practice. Chicago: JAMA & Archives Journals AMA 2001.
3. Carneiro AV. Níveis de evidência científica e graus de recomendação de esquemas de tratamento. Informação Terapêutica 2000;Jan-Mar:03-04.
4. Heidenreich PA, McDonald KM, Hastie T, Fadel B, Hagan V, Lee BK, Hlatky MA. Meta-analysis of trials comparing b-blockers, calcium antagonists, and nitrates for stable angina. JAMA 1999;281:1927-36.
5. Guyatt GH, Sackett DL, Cook DJ, The Evidence-Based Medicine Working Group. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? JAMA 1994;271:59-63.
6. Evidence-Based Medicine. How to practice and teach EBM. Edinburgh: Churchill Livingstone 2000.
7. Garg R, Yusuf S. Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. JAMA 1995;273:1450-6.
8. Waldo AL, Camm AJ, deRuyter H, for the SWORD Investigators. Effect of d-sotalol on mortality in patients with severe left ventricular dysfunction after recent and remote myocardial infarction: survival with oral d-sotalol. Lancet 1996;348:7-12.
9. Greene WL, Concato J, Feinstein AR. Claims of equivalence in medical research: are they supported by the evidence? Ann Int Med 2000;132:715-22.
10. LaRosa JC, He J, Vupputuri S. Effect of statins on risk of coronary disease. JAMA 1999;282:2340-6.
11. West of Scotland Coronary Prevention Study Group. Influence of pravastatin and plasma lipids on clinical events in the West of Scotland Coronary Prevention Study (WOSCOPS). Circulation 1998;97:1440-5.

---

# XXIV

## Congresso Português de Cardiologia

FUNCHAL – 26 a 30 de Abril de 2003

Organiza<sup>ção</sup>

**Direcção da Sociedade Portuguesa de Cardiologia**

---

*Secretariado*

SOCIEDADE PORTUGUESA DE CARDIOLOGIA

Campo Grande, 28, 13.º

1700-093 LISBOA

Telefs. 21 781 7634 / 30 – Fax 21 793 1095

e-mail: congresso@mail.spc.pt